**S2 Table. Summary statistics for the *de novo* transcriptome assembly of Illumina reads from the dioecious line 03-009 obtained using Trinity.**

|  |  |  |  |
| --- | --- | --- | --- |
| 　 | Original contigs | Non-redundant contigsa | Non-redundant unigenesb |
| Number of contigs | 249,622 | 190,200 | 189,948 |
| Total length of contigs | 133,464,632 | 66,816,943 | 66,030,647 |
| Average length of contigs | 534 | 351 | 347 |
| Maximum length of contigs | 34,519 | 34,073 | 15,770 |
| Minimum length of contigs | 101 | 101 | 101 |
| N50 length | 1,474 | 607 | 591 |

aNon-redundant contigs were selected from the original contigs using RSEM 1.2.3.

bNon-redundant unigenes were generated by removal of contaminating sequences (using the spinach chloroplast genome [AJ400848.1, NC\_002202.1], *A. thaliana* mitochondrial genome [NC\_001284.2], *A. thaliana* RNAs from NCBI [741 rRNAs, 1690 tRNAs, 792 ncRNAs, 555 transcribed RNAs, 48 other RNAs, and 4 genomic RNAs], human genome [GRCh37], bacterial genomes [NCBI RefSeq], UniVec, fungal genomes [NCBI RefSeq], and PhiX174 [NC\_001422.1]) from the non-redundant contigs.